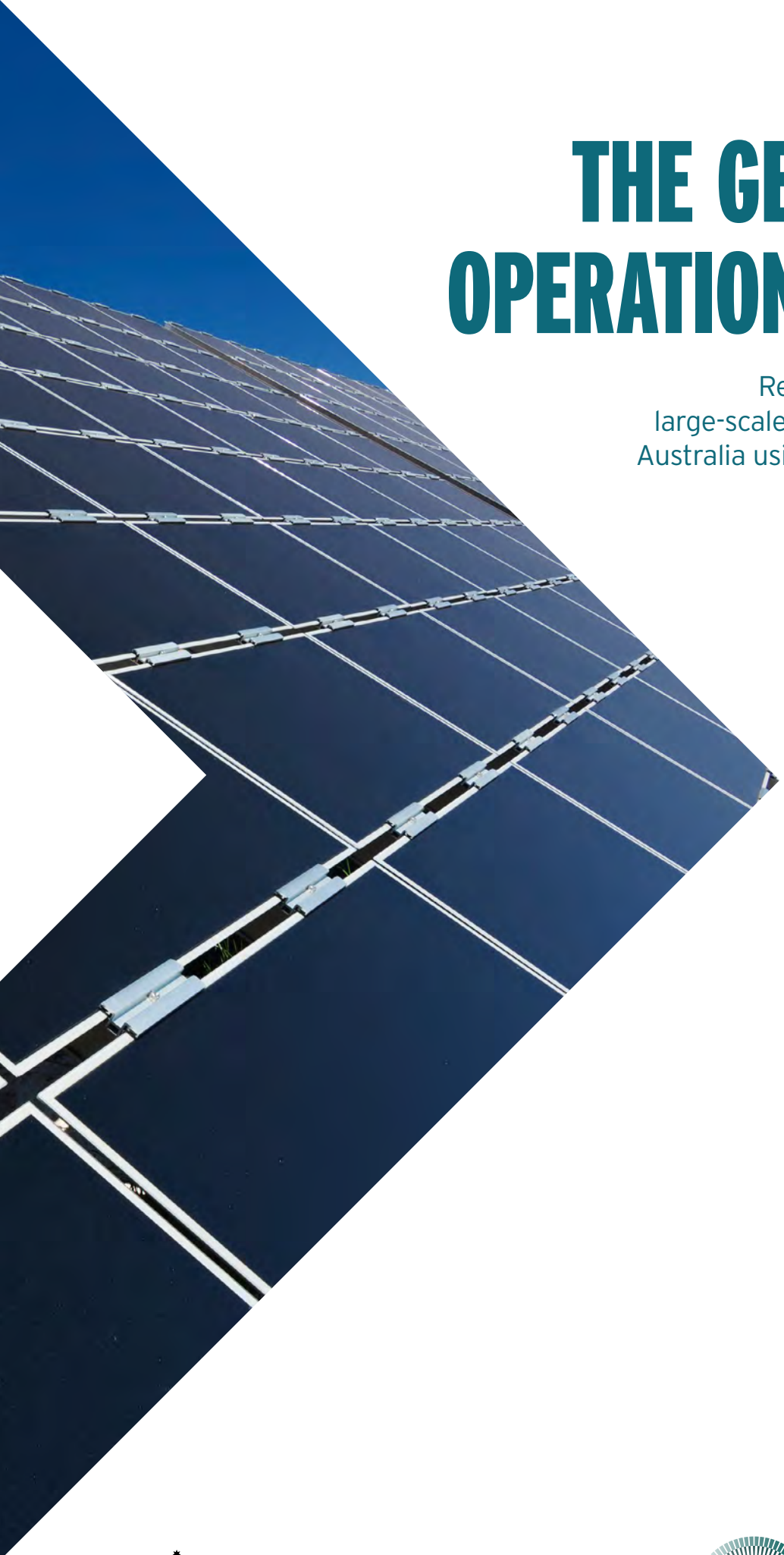

# THE GENERATOR OPERATIONS SERIES

## Report Five: Benchmarking large-scale solar PV performance in Australia using satellite weather data

**AUGUST 2022**

Australian Government
Australian Renewable Energy Agency
ARENA
cefc
Ekistica

# EXECUTIVE SUMMARY

Solar photovoltaic (PV) performance depends on system quality and weather and can be quantified using the performance ratio (PR) calculation. PR is the ratio of the electricity generated to the electricity that would have been generated if the plant consistently converted sunlight to electricity at the level expected from the DC nameplate rating [1] (see Definitions).

PR measurements at solar farms typically use high accuracy ground station irradiance and temperature measurements to calculate the performance of a solar farm given the observed weather conditions. The reliance on privately-owned high-quality ground station data means that; performance data for solar farms is temporally and spatially scarce, is expensive to maintain, and remains unavailable to the public. Without this data, it is impossible to compare performance between projects or benchmark the performance of solar farms operating in a particular market.

However, the accuracy of satellite data has significantly improved in recent years, to the extent that it may be possible to use satellite weather data and public electricity generation data to calculate PRs for large scale solar farms across the NEM. This would enable the benchmarking of performance across solar farms. This study quantifies the level of accuracy that satellite-based performance measurements can achieve and uses this result to estimate the average and distribution of performance for solar farms operating on the NEM in 2020.

As part of ARENA's knowledge sharing, funding recipients are required to submit confidential project information, such as ground station weather data. This study compares the irradiance, PV cell temperature and PR estimates based on satellite data with estimates based on high accuracy ground station data from the 10 projects in ARENA's LSS Funding Round. This is used to quantify the accuracy and error of PR values calculated using satellite data. The analysis is then expanded to calculate PRs for 26 solar farms across the NEM. The mean PR across all farms in the study is calculated, with confidence intervals quantified using the error calculated using LSS project data.

While it is preferable to use ground-based weather data to perform PR calculations on large-scale solar farms, this analysis shows that it is, within reason and depending on the application, possible to rely on satellite data, albeit introducing some additional uncertainty. This finding is especially useful in circumstances where ground-based weather data is not available or where the performance of multiple farms is being compared or assessed in aggregate. Based on the analysis of irradiance, cell temperature and performance ratios across 10 of ARENA's LSS sites in the year 2020,[1] satellite data was found to align closely with ground station measurements, achieving the following levels of accuracy to a 95 per cent confidence level:

- Annual irradiance estimates are within ± 4.6 per cent of ground station measurements.
- 5-minute cell temperature estimates are within ± 8.0 per cent of ground station measurements. However, uncertainty from satellite-based cell temperature estimates introduces additional uncertainty of ± 1.1 per cent to annual temperature-corrected performance ratio calculations.
- Temperature-corrected performance ratios are within ± 4.8 per cent of ground station measurements. This level of accuracy reflects both the observed error in the performance ratios across the LSS sites and the uncertainties of the input irradiance and cell temperature data as shown.
- This study provides benchmarks of performance across the NEM at 26 solar farms in 2020, including error estimates, using satellite data. The key results are as follows:
- The average raw performance ratio was 68.4 per cent, with the middle 50 per cent of farms in the sample falling within the range of 63.3 per cent to 74.3 per cent. This indicates overall performance, including the impacts of curtailment.
- The average unconstrained performance ratio was 75.4 per cent, with the middle 50 per cent of farms in the sample falling within the range of 71.2 per cent to 79.7 per cent. This indicates the performance of farms excluding periods of curtailment and other constraints on farm output.
- The average temperature-corrected performance ratio was 79.3 per cent, with the middle 50 per cent of farms in the sample falling within the range of 75.4 per cent to 83.5 per cent. These figures are corrected for losses due to module cell temperature, which enables comparison of farms in different regions and better estimates of relative performance at farms with several weeks of excluded data.

1 Note that estimations of irradiance and cell temperature are from pvlib [7] (see Definitions) using satellite weather data.

The range of performance for the middle 50 per cent of farms in the sample includes uncertainty due to satellite estimates and is provided at a 95 per cent confidence level. This data can be used to determine whether a given solar farm's performance is in the top or bottom 25 per cent of performance, relative to those sampled. The mean temperature-corrected performance ratio at farms in Queensland, NSW and Victoria were similar (78.3 per cent to 79.8 per cent). South Australian (SA) farms exhibited a substantially higher mean performance (83.0 per cent), though this was based on a sample size of 2. The mean performance ratio improved by over 10 per cent for farms in Victoria and SA, compared to less than 5 per cent for the average in Queensland and NSW. This likely reflects both a higher level of curtailment occurring in the former states and slightly smaller sample sizes.

AVERAGE PERFORMANCE INCLUDING UNCERTAINTY FROM SATELLITE MEASUREMENTS

| | RAW PR [%] | UNCONSTRAINED PR [%] | TEMPERATURE CORRECTED PR [%] |
|---|---|---|---|
| **Mean** | 68.4 | 75.4 | 79.3 |
| **PR Range for Mean (95% Confidence)** | 67.7 - 69.0 | 74.7 - 76.1 | 78.6 - 80.1 |
| **PR Range for Middle 50% of Solar Farms (95% Confidence)** | 63.3 - 74.3 | 71.2 - 79.7 | 75.4 - 83.5 |

Above image: Broken Hill Solar Plant

# CONTENTS



Cover image: Rottnest Solar Farm

# ABBREVIATIONS

| | |
|---|---|
| AEMO | Australian Energy Market Operator |
| ARENA | Australian Renewable Energy Agency |
| CEFC | Clean Energy Finance Corporation |
| DHI | Direct Horizontal Irradiance |
| DNI | Direct Normal Irradiance |
| EPC | Engineering, Procurement and Construction |
| GHI | Global Horizontal Irradiance |
| LSS | Large-scale Solar |
| NEM | National Energy Market |
| O&M | Operation and Maintenance |
| POA | Plane of Array |
| PPA | Power Purchase Agreement |
| PR | Performance Ratio |
| PV | Photovoltaic |

## PERFORMANCE RATIOS UNDERSTOOD

This study uses satellite data to measure the performance of solar farms on the NEM. Performance is measured by calculating the performance ratio for each site.

Performance ratios are a crucial metric that allows investors, developers and operators of solar farms to track the performance of the asset. Liquidated damages may be payable under EPC contracts, O&M contracts or PPAs where solar farms do not achieve performance ratios exceeding a pre-agreed value. Consistently high performance ratios relative to those predicted by yield models from the time of financial close can be used to demonstrate the reliability of an asset as part of refinance transactions for solar farms or renewable energy portfolios.

The performance ratio calculation requires accurate measurements of generation, irradiance falling on modules, and PV module temperature as inputs. Generation data is publicly available at 5-minute resolution from AEMO. The irradiance falling on the modules and PV module temperature can be estimated using satellite data, provided reasonable assumptions can be made about the setup of the solar farm. This paper begins by validating these assumptions and quantifying the accuracy of the irradiance and PV cell temperature estimates thus derived from satellite data, by comparing satellite-based estimates with estimates from ground station data measured at 10 projects in ARENA's LSS portfolio. The raw, unconstrained and temperature-corrected performance ratios are then calculated for these 10 projects and the uncertainty of these measurements is quantified. Finally, raw, unconstrained and temperature-corrected performance ratios were calculated for all 26 solar farms on the NEM for which sufficient data was available in 2020.

## IRRADIANCE

Solar irradiance is the amount of solar energy that arrives at a specific area of a surface during a specific time interval, typically measured in Watts per square metre [2]. A fundamental step in calculating solar PV performance (see Figure 1) is knowing how much irradiance is reaching the solar array in its plane-of-array (POA). Calculating this also requires knowledge of the installation set-up at a solar farm so that the orientation of the POA with respect to the sun can be determined for any given time interval. This includes knowing whether a site is a single-axis tracking or fixed-tilt array, its orientation, azimuth etc. If a site is fixed-tilt, the angle at which it is tilted also needs to be known.

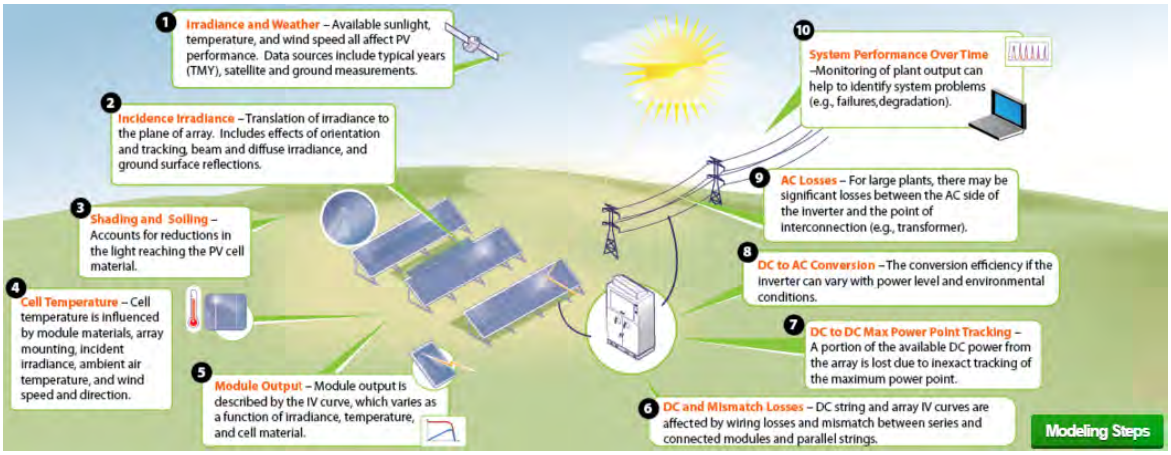


Figure 1. A list of steps that can be taken to calculate performance of solar PV systems. Image sourced from [3].

After extensive research, no known public database (i.e., free access) providing installation details of solar farms in Australia seems to exist. As a result, this study attempts to automatically determine the installation set-up at each solar farm by modelling a range of different set-ups and comparing the generation profiles to the Australian Energy Market Operator's (AEMO) public dispatch data.

POA irradiance can be calculated using various components of solar irradiance including Global Horizontal Irradiance (GHI), Direct Horizontal Irradiance (DHI) and Direct Normal Irradiance (DNI). Estimating POA irradiance requires a sequence of models including the decomposition and transposition models. Transposition is the calculation of the incident irradiance on a tilted plane, from the horizontal irradiance data [4].

Estimations of POA irradiance may differ depending on the transposition model chosen, thus affecting PR calculations. It is therefore important to have the correct transposition model so financial metrics and evaluations of solar farms can be accurately determined by investors and developers. The transposition model used in pvlib is, by default, the Perez model. The Perez model is a more sophisticated[2] model when well measured GHI data is available [4]. More details on the Perez model are provided in Definitions.

Satellite irradiance estimates used in this analysis are from Solcast [5]. Solcast estimate irradiance in four key steps, which are simplified below (more details on this can be found in [6]):

1. Processes satellite raw imagery through a range of algorithms.
2. Using sophisticated statistical models, a representation of the 'background', more commonly referred to as the 'albedo', is generated and allows distinction to be made between cloud cover, snow, bright sand, ocean glare or other aspects of the imagery that could be incorrectly interpreted as cloud.
3. Decomposition of satellite imagery into cloudy and cloud-free regions.
4. A clear-sky radiation model allows for global aerosol (dust, salt, smoke, etc.) and water vapor content to generate precise estimates of the solar radiation available to cloud-free regions. For areas with cloud cover, the cloud opacity is used to produce estimates of the total amount of solar irradiance reaching the Earth's surface.

2 Perez is more sophisticated than the Hay and isotropic transposition models in that it takes into account both the circumsolar and the horizon diffuse irradiance.

The POA irradiance calculated from satellite-derived global GHI, DHI and DNI using the decomposition and transposition models represents the irradiance reaching the surface of the module. However, module performance is a function of irradiance reaching the cells within the module glass casing. Therefore, reflective and other irradiance losses must be accounted for when calculating the performance of a solar PV system.

The POA irradiance with these losses accounted for is referred to as effective irradiance, calculated here using pvlib [7]. This study uses effective irradiance instead of global tilted irradiance to quantify the irradiance used for performance ratio calculations. While effective irradiance values are slightly lower than global tilted irradiance values due to the losses mentioned above, and the latter are typically used for PR calculations and are closer to pyranometer readings, in practice, pyranometer readings are often taken as equivalent to effective irradiance [8].

## DETECTING THE SET-UP AT EACH SOLAR PV SITE

### ESTIMATING AZIMUTH

Australia being in the southern hemisphere results in most Australian solar farms facing true north, or very close to it, as this optimises total irradiance falling on the modules over the course of the year. However, this is not always possible due to physical constraints such as limited land area and non-uniform land boundaries. Orienting an array axis away from north may trade system performance with ease of installation or increased array capacity within the project area [9]. Sometimes physical constraints, such as land boundaries, result in an azimuth other than true north being preferable [9]. Global Sustainable Energy Solutions (a global PV installation and consultancy specialist) suggest that the azimuth can vary up to 15 degrees either way without significant losses (e.g., 27 – 31 kWh/kWp/yr, equivalent to approx. 2 per cent of annual output) [9]. As a result, we have simplified the analysis by assuming azimuth is facing true north for all solar PV systems (see Figure 2). However, this assumption will sightly degrade the accuracy of the irradiance and cell temperature estimations from satellite weather data.

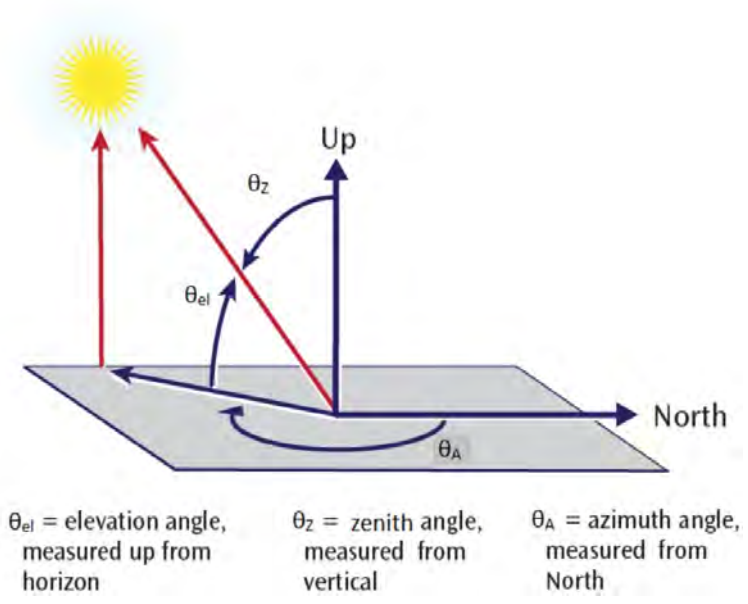


Figure 2. The position of the sun relative to an observer on the surface of the Earth. Image sourced from [10].

## SINGLE-AXIS TRACKING OR FIXED-TILT ARRAYS?

Single-axis tracking solar farms will generally follow the same tracking algorithm that maximises the POA irradiance under clear sky conditions, however individual solar farms may implement slightly different tracking algorithms, such as backtracking[3] or dynamic algorithms that respond to locally measured irradiance conditions (e.g. Nextracker's TrueCapture technology). Since optimisations of the latter kind are a more recent development, it is unlikely they were already deployed on most solar farms in 2020, when the data for this analysis was collected. The single-axis tracking simulations undertaken in this study assume backtracking algorithms are employed[4] and an array's maximum array tilt angle is 60° (see Figure 3 [11]).

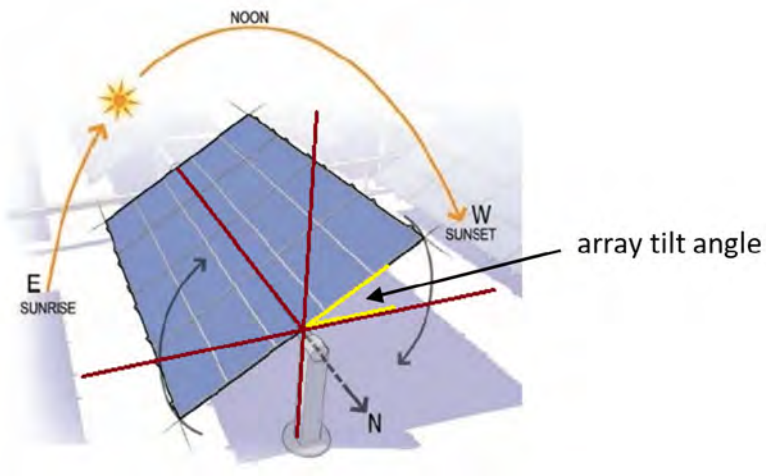


Figure 3. Schematic of a single-axis solar PV tracking array showing the array tilt angle between the yellow lines. Original image from [11].

A total of 52 pvlib simulations were conducted for each site, consisting of one with single-axis tracking and fifty with a fixed-tilt iterating over each degree from zero to 50 degrees. The simulations at each site were then compared to the AEMO's publicly recorded dispatch data. The comparison that had the strongest correlation with AEMO dispatch data (i.e., highest coefficient of determination, or 'r-squared') indicated which simulation most likely resembled the site's set-up.

Prior to conducting this analysis, periods of data where the solar farm was or was likely to be operating at a sub-optimal level were excluded. This is so that the solar farm set-up detected through this method would be matched to the array under normal operating conditions. The following periods of data were excluded from this analysis:

1. Night-time periods: times when the clear-sky global horizontal irradiance was equal to zero.
2. Extreme nominal generation: where generation is below 10 per cent or above 90 per cent of AC capacity.
3. Negative pricing: where the 5-minute spot price or 30-minute settlement price were negative.
4. Semi-dispatch cap: where the semi-dispatch cap was active requiring the solar farm to follow dispatch signals.
5. Local limit active at solar farm (this was manually estimated by visualising data)
    a. AEMO have begun recording this data publicly since November 2021.
    b. Entire days were omitted where a local limit was identified to have applied at any point during the day.

3 The behaviour of arrays operating at a lower surface tilt when the sun is low in the sky to prevent row-to-row shading is known as "backtracking". It is accomplished by the tracker rotating backwards from the "ideal" rotation so that the row's shadow is shortened and misses the row behind it [29].

4 Preliminary analysis showed that assuming backtracking was employed reduced the average mean bias error between satellite and ground plane-of-array irradiance at LSS sites.

These exclusions are illustrated in Figure 4 below.

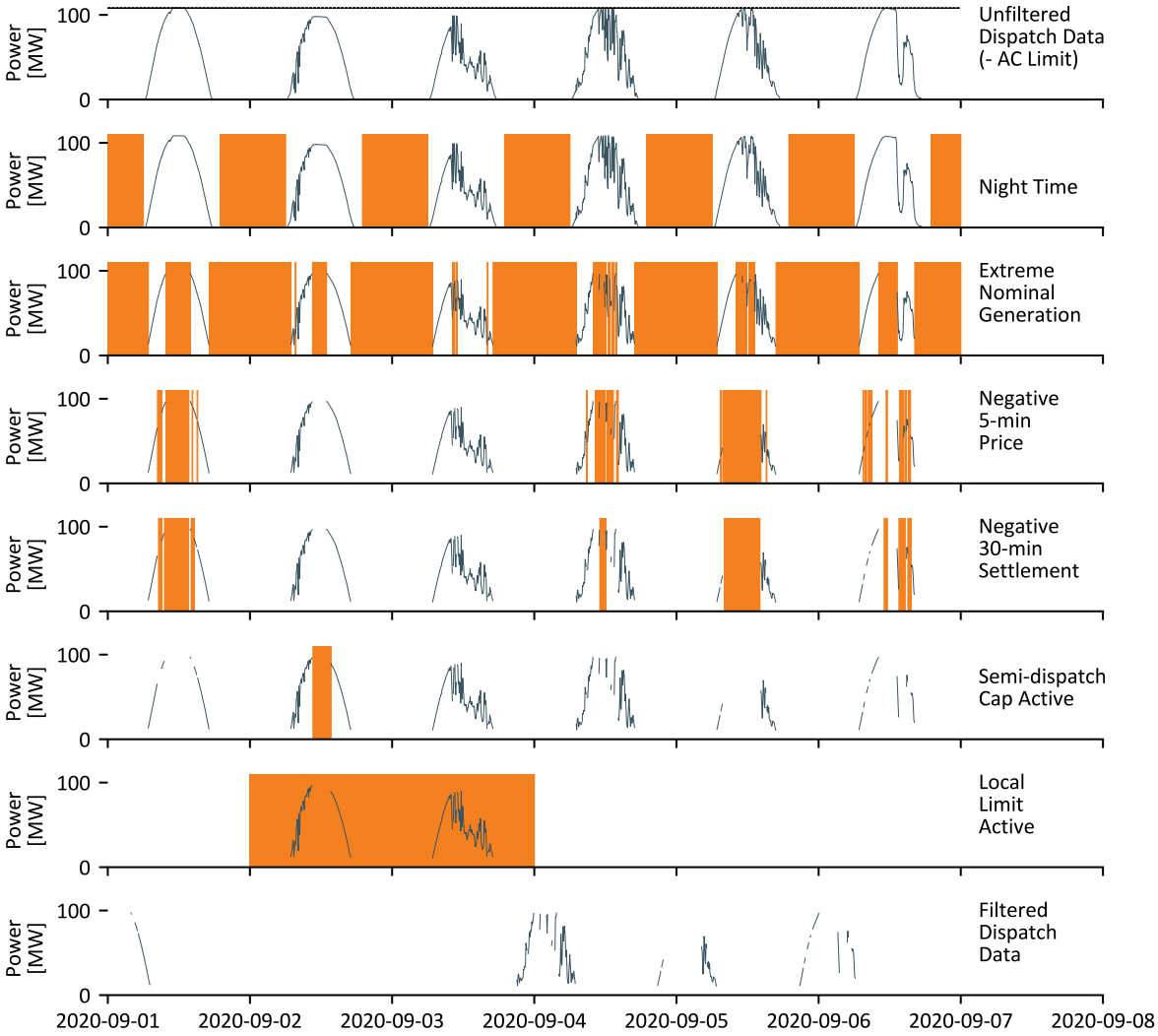


Figure 4. Visualising the process of filtering out periods from AEMO's NEMWEB dispatch data not suitable for performance ratio analysis

After excluding unsatisfactory periods of data in 2020, twenty-six solar farms had sufficient data remaining to be included in the remainder of the analysis.

The likely set-up at each site is described in Table 1, including the amount of training data available.[5] While some sites have more limited training data, the remaining data should still have a stronger correlation with the true array set-up than with other configurations. However, sites with limited amounts of training data may have less reliable results for set-up and subsequent analysis. In particular, local limits (e.g., inverters offline, curtailment instructions from AEMO provided externally to the dispatch engine) have been applied conservatively, with the entire day being excluded if a local limit was observed to have applied (by visual inspection of the data) at any point during the day. This limits the amount of data available for detecting the likely set-up and for calculating the unconstrained and temperature-corrected performance ratios. This stage of data cleansing is particularly difficult because local limit data tags in 2020 are not publicly available through AEMO. AEMO began publicly recording local limit data in November 2021 and therefore this data cleansing stage will be more accurate and quicker to implement on operational data post November 2021.

5 Training data is the data used at each site to determine the likely set-up. The sample size of training data changes site-to-site as each site experiences curtailment differently, resulting in different periods of unsatisfactory data requiring to be removed prior to analysis. The number of 'Hours' are calculated as the total number of 5-minute intervals divided by 12. The number of days 'Days' are counted as dates for which there were at least one hour's worth of measurements after exclusions were applied.

TABLE 1. LIKELY SET-UP DETERMINED AT EACH SOLAR PV SITE

| SITE | LIKELY SET-UP | HOURS OF TRAINING DATA | DAYS OF TRAINING DATA | STATE |
|---|---|---|---|---|
| Coleambally | Single-Axis Tracking | 594.8 | 106 | NSW |
| Finley | Single-Axis Tracking | 1194.2 | 183 | NSW |
| Gullen Range | Fixed-Tilt 29° | 1929.6 | 288 | NSW |
| Manildra | Single-Axis Tracking | 1987.4 | 264 | NSW |
| Moree | Single-Axis Tracking | 2121.5 | 278 | NSW |
| Nyngan | Fixed-Tilt 23° | 2532.3 | 366 | NSW |
| Parkes | Single-Axis Tracking | 2188.2 | 316 | NSW |
| White Rock | Fixed-Tilt 24° | 2649.2 | 356 | NSW |
| Clermont | Single-Axis Tracking | 2576.2 | 364 | QLD |
| Collinsville | Fixed-Tilt 13° | 2203.8 | 365 | QLD |
| Darling Downs | Fixed-Tilt 24° | 2072.9 | 299 | QLD |
| Daydream | Single-Axis Tracking | 2060.7 | 272 | QLD |
| Emerald | Single-Axis Tracking | 2544.9 | 363 | QLD |
| Hamilton | Single-Axis Tracking | 485.2 | 75 | QLD |
| Kidston | Single-Axis Tracking | 998.7 | 216 | QLD |
| Oakey 1 | Single-Axis Tracking | 979 | 153 | QLD |
| Ross River | Single-Axis Tracking | 2290.5 | 353 | QLD |
| Rugby Run | Single-Axis Tracking | 387.4 | 70 | QLD |
| Whitsunday | Single-Axis Tracking | 219.3 | 37 | QLD |
| Bungala One | Single-Axis Tracking | 1675.5 | 311 | SA |
| Tailem Bend 1 | Fixed-Tilt 19° | 1889.5 | 322 | SA |
| Bannerton | Single-Axis Tracking | 233.6 | 58 | VIC |
| Gannawarra | Single-Axis Tracking | 1053.7 | 143 | VIC |
| Karadoc | Single-Axis Tracking | 1172.6 | 185 | VIC |
| Numurkah | Single-Axis Tracking | 2250.6 | 348 | VIC |
| Wemen | Single-Axis Tracking | 1309.5 | 218 | VIC |

## COMPARING IRRADIANCE FROM GROUND-BASED MEASUREMENTS WITH ESTIMATES DERIVED FROM SATELLITE WEATHER DATA

Figure 5 compares POA irradiance from ground weather station data (henceforth referred to as "site irradiance") and estimates of effective irradiance derived from Solcast[6] satellite weather data (henceforth referred to as "satellite irradiance") on three of ARENA's LSS projects[7]. The difference in irradiance profiles indicates that Darling Downs is a fixed-tilt array, while Gannawarra and Parkes are both single-axis tracking arrays, as predicted by the method described in the previous section.

Figure 6 demonstrates a reasonable level of accuracy of satellite irradiance with respect to site irradiance. Note the ability for ground weather stations to more effectively capture the impact cloud cover has on blocking effective irradiance, marked by the sharp rises and drops in ground weather station data.

6 For more information on how Solcast estimates irradiance, see [6].

7 Gannawarra did not receive funding under ARENA's LSS program but contributes data for Knowledge Sharing.

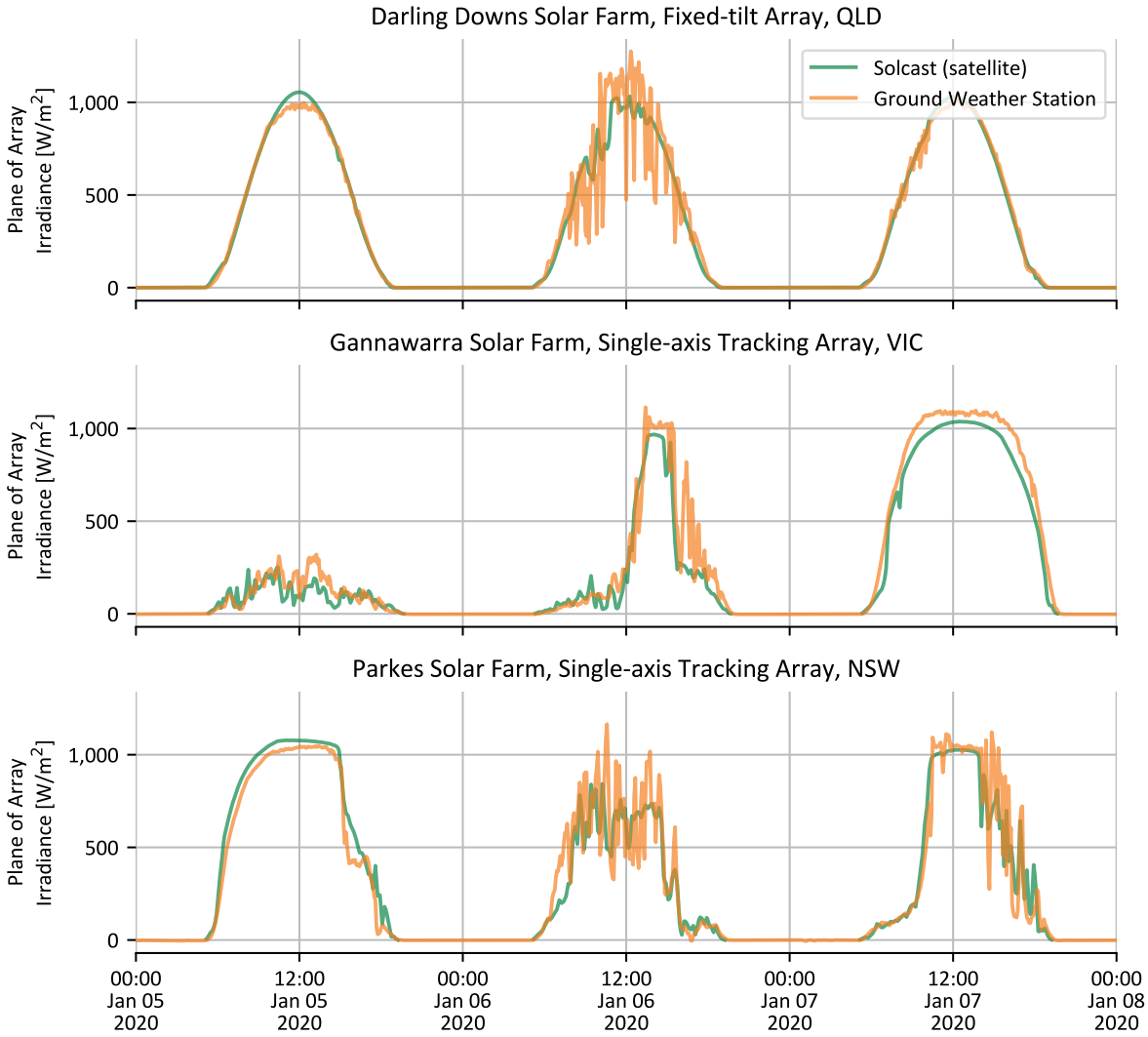


Figure 5. Comparing effective irradiance from ground weather station and satellite data across three of ARENA's LSS projects over three individual days.

## QUANTIFYING UNCERTAINTY

Quantifying the uncertainties associated with models of solar farm performance is a complicated process that is the subject of ongoing research within the industry [12, 13], but it is important to enable investors and developers to better understand the short-term and long-term risks associated with a project. This paper reports on the uncertainty associated with satellite estimates of the inputs and outputs of PR calculations using the error statistics adopted by the National Renewable Energy Laboratory[8] (NREL) to evaluate modeled data in [13]: bias, standard deviation of bias, root mean square error (RMSE), and standard and expanded uncertainty.

Definitions on the key error statistics and how they are calculated can be found in Definitions. In simple terms, the commonly referred to error statistics are described below:

- Mean Bias Error (MBE): Indicates whether a model consistently overestimates or underestimates the target value.
- Standard Deviation of Bias (SD): Indicates the variation in the bias over the modelled values. Where bias estimates are available for a number of sites for a given model, the standard deviation of bias can be used to estimate the uncertainty to be expected in measurements obtained by applying the model to a new site [14, 15].
- Root Mean Square Error (RMSE): Measures the average error of the model without considering error direction and gives a relatively high weight to large errors. Reflects both the bias and variation in the error of the data.
- Standard Uncertainty (u): The standard uncertainty associated with an estimate of a quantity, equal to one standard deviation. Standard uncertainty values can be estimated directly from statistical methods applied to test data (Type A) or by other means, such as 'scientific judgment, experience, specifications, comparisons, and calibration data' (Type B) [13]. Uncertainty values provided by equipment manufacturers for pyranometers and other equipment are an example of Type B standard uncertainties.
- Expanded Uncertainty (U): The uncertainty associated with an estimate at a given confidence level. The standard uncertainty can be converted to an expanded uncertainty by multiplying by a coverage factor (k). Where the uncertainty follows a normal distribution, coverage factor is approximately k = 2 for a 95 per cent confidence level.

Each of these statistics can be expressed as absolute values or as relative errors expressed as a percentage of the mean value of the quantity being estimated. Other than performance ratios,[9] errors presented as percentages throughout this paper are the absolute error as a percentage of the mean of the value being measured.

The error in the satellite irradiance estimates compared to the site irradiance measurements are summarized in Figure 6 and Table 2 below. The overall spread in the error distribution reduces as the temporal resolution reduces from 5-minutes to monthly estimates, as illustrated by the reduced range and interquartile range shown on the boxplots and the reduction in RMSE. This is because the random errors associated with individual measurements at short time scales cancel out as the measurement is summed over longer time scales, and is consistent with observations from previous studies comparing satellite and ground station measurements [16].[10]

Overall, the average MBE across the different sites at both the 5-minute and monthly time resolutions is less than 0.25 per cent. At none of the sites does the MBE exceed a magnitude of 4 per cent. This indicates that the data can be relied upon for the calculation of performance ratios with a relatively high level of accuracy.

8 NREL is a national laboratory of the U.S. Department of Energy.

9 For performance ratios, errors expressed as percentages are the absolute error in the performance ratio estimate, as performance ratios themselves are expressed as percentages. Where errors in performance ratio measurements are expressed as a proportion of the value being estimated, the term 'relative' error is used.

10 Note that the monthly irradiance values are measured in $Wh/m^2$, as the irradiance has been summed over time, whereas the 5-minute values are reported in $W/m^2$, indicating instantaneous measurements.

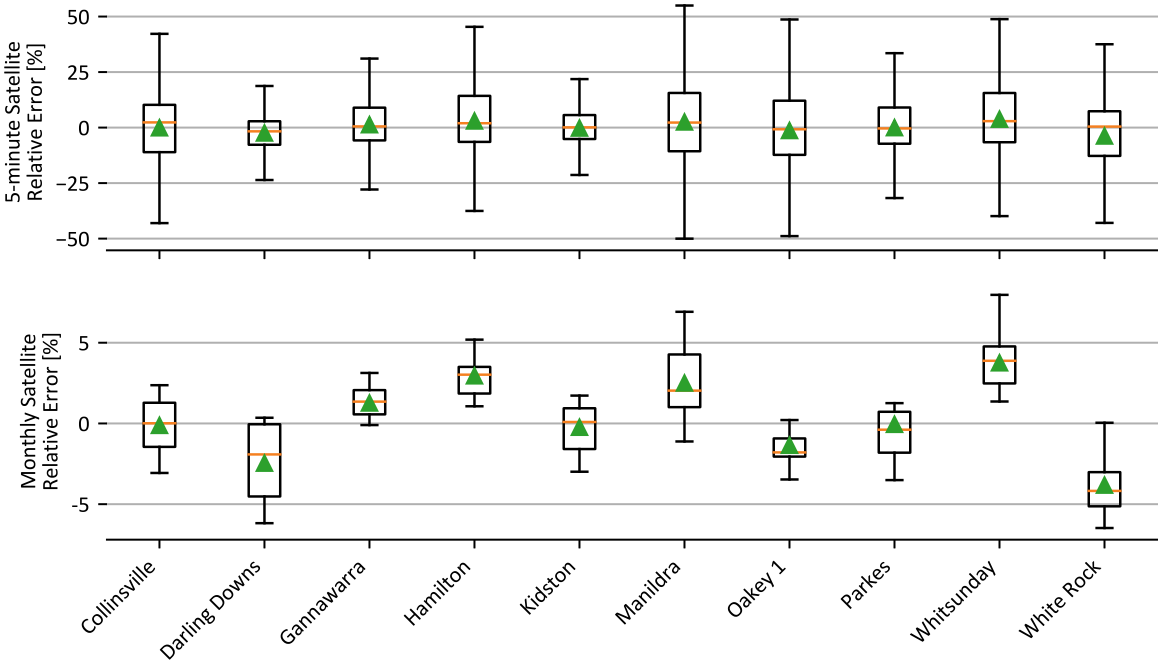


Figure 6. Describing the error between satellite and ground weather station POA irradiance at ten of ARENA's LSS projects (mean = green triangle, median = orange line). Error shown is as a percentage of the mean site value for the given time resolution.

TABLE 2. MBE AND RMSE FOR FIVE-MINUTE AND MONTHLY IRRADIANCE ESTIMATES AT EACH PROJECT

| PROJECT | 5-MINUTE | | | | MONTHLY | | | |
|---|---|---|---|---|---|---|---|---|
| | MBE | | RMSE | | MBE | | RMSE | |
| | [W/M²] | [%] | [W/M²] | [%] | [WH/M²] | [%] | [WH/M²] | [%] |
| Collinsville | -0.6 | -0.12 | 115.7 | 24.01 | -213 | -0.12 | 2951 | 1.71 |
| Darling Downs | -12.1 | -2.46 | 111.4 | 22.73 | -2473 | -2.46 | 3547 | 3.53 |
| Gannawarra | 6.9 | 1.26 | 134.1 | 24.71 | 2516 | 1.26 | 3797 | 1.91 |
| Hamilton | 17.1 | 2.95 | 166.9 | 28.68 | 6282 | 2.95 | 8178 | 3.84 |
| Kidston | -1.6 | -0.24 | 138.0 | 21.17 | -576 | -0.24 | 3413 | 1.42 |
| Manildra | 13.0 | 2.50 | 194.8 | 37.45 | 4782 | 2.50 | 6604 | 3.45 |
| Oakey 1 | -7.9 | -1.36 | 210.3 | 36.14 | -2902 | -1.36 | 4132 | 1.93 |
| Parkes | -0.3 | -0.06 | 144.6 | 26.46 | -119 | -0.06 | 5366 | 2.67 |
| Whitsunday | 21.7 | 3.75 | 176.5 | 30.54 | 7881 | 3.75 | 9372 | 4.46 |
| White Rock | -18.7 | -3.82 | 125.3 | 25.66 | -6393 | -3.82 | 7117 | 4.26 |
| **Mean** | 1.8 | 0.24 | 151.7 | 27.76 | 879 | 0.24 | 5448 | 2.92 |
| **Std. Dev.** | 12.3 | 2.29 | 32.1 | 5.21 | 4229 | 2.29 | 2131 | 1.07 |

The standard deviation of bias across all sites can be used to quantify the standard and expanded uncertainty for estimates of total monthly or annual irradiance.[11] Total annual irradiance is the major source of uncertainty in performance ratio calculations. A standard uncertainty of approx. ±2.3 per cent can be adopted.[12] Monthly or annual totals at other sites can be expected to be within ±4.6 per cent of ground station measurements with 95 per cent confidence. A similar value could be adopted for 5-minute measurement given the similarity in mean and standard deviation of MBE values at the 5-minute and monthly levels of aggregation. The total uncertainty in the estimate can be obtained by combining the uncertainty relative to ground station measurements with the inherent uncertainty in the ground station measurements, by summing the uncertainty values in quadrature. Secondary standard pyranometers have an expanded uncertainty of ±2 per cent at the 95 per cent confidence level for daily totals, and this can be adopted as a conservative estimate for the uncertainty in monthly or annual totals [14].

11 The bias values for each site were identical at monthly and annual aggregation.

12 Symmetrical uncertainties are typically adopted provided that the MBE is close to zero [15].

The resulting total expanded uncertainty in the satellite-based monthly or annual irradiance is ±5.0 per cent at the 95 per cent confidence level. This means that there is a high degree of confidence that monthly or annual totals at other sites will fall within ±5.0 per cent of the true value.

TABLE 3. UNCERTAINTY FOR SATELLITE-BASED MONTHLY AND ANNUAL IRRADIANCE ESTIMATES

| | UNCERTAINTY RELATIVE TO GROUND STATION MEASUREMENTS | TOTAL UNCERTAINTY INCLUDING GROUND STATION INSTRUMENT UNCERTAINTY |
|---|---|---|
| **Standard Uncertainty** | ±2.3% | ±2.5% |
| **Expanded Uncertainty at 95% confidence level** | ±4.6% | ±5.0% |

Image: Oakey Solar Farm

# TEMPERATURE

Given a set irradiance, solar PV generation reduces the further cell temperatures exceed standard testing conditions (i.e., 25°C). Put simply, PV performance decreases as ambient temperature increases, this is true of all electronic devices and systems. Therefore, two identical solar farms receiving the same solar resource will generate different amounts of energy if the solar modules are at different temperatures. However, the underlying performance of the solar farms may be identical, apart from the difference in losses due to module temperature, resulting in different generation. Solar PV performance calculations can account and correct for temperature associated electricity generation losses. If temperature is not corrected for in performance ratio calculations, then solar PV appears to perform worse in hotter months. These corrections are particularly important where farms in different locations are being compared or where limited data is available.

## DETECTING CELL TEMPERATURE AT EACH SITE

Actual cell temperature data has been collected from ARENA's LSS projects where it is measured using equipment directly on site. Comparisons to these actual cell temperature recordings are made against estimations of cell temperature based on satellite weather data. pvlib [7] estimates cell temperature from satellite weather data and these data are henceforth referred to as satellite cell temperature. The pvlib cell temperature model employed in this study is an empirically-based thermal model developed by Sandia National Laboratories [17].

It is important to draw the distinction between PV cell temperature and PV module back surface temperature. Firstly, the model takes, from the satellite-derived data, the solar irradiance on the PV module front surface, ambient air temperature and wind speed at standard 10 m height along with two empirically derived coefficients to determine the temperature of the module's back-surface [2].

These coefficients, 'a' and 'b', establish the upper limit for module temperature at low wind speeds and high solar irradiance, and the rate at which module temperature drops as wind speed increases, respectively. The cell temperature is then derived from the module back-surface temperature by adopting a one-dimensional thermal conduction model through the materials between the back surface and the cell.

The ratio between actual solar irradiance on the module and reference solar irradiance on module (1,000 W/m$^2$) is used alongside the empirically-derived temperature differences between the cell and the back surface at reference condition, which is typically in the range of 2 to 3 degrees Celsius for flat-plate modules in an open rack mount.

This study assumes that solar PV modules have a glass/cell/polymer sheet construction with an open rack mount and uses the recommended temperature module parameters in Table 4 from Sandia's PV Array Performance Model (SAPM) [18]. It is common practice to fine tune the thermal model by determining new values for these coefficients to adapt to the site conditions [2]. The 'a' coefficient was iterated over a range of values until the MBE in cell temperature between satellite-derived data and ground station weather data was minimised. A lower value of 'a' suggests that, on average, the modules deployed in ARENA's LSS projects have a higher upper limit for module temperature at low wind speeds and high solar irradiance.

TABLE 4. TEMPERATURE MODULE PARAMETERS

| | A | B | ΔT (°C) |
|---|---|---|---|
| **SAPM Recommended** | -3.56 | -0.075 | 3.0 |
| **Adapted** | -4.1 | -0.075 | 3.0 |

## COMPARING CELL TEMPERATURE FROM GROUND-BASED MEASUREMENTS WITH ESTIMATES DERIVED FROM SATELLITE WEATHER DATA

Figure 7 compares, across three days, site and satellite cell temperature recordings at three of ARENA's LSS generators. The plots indicate relatively close alignment between the two datasets.

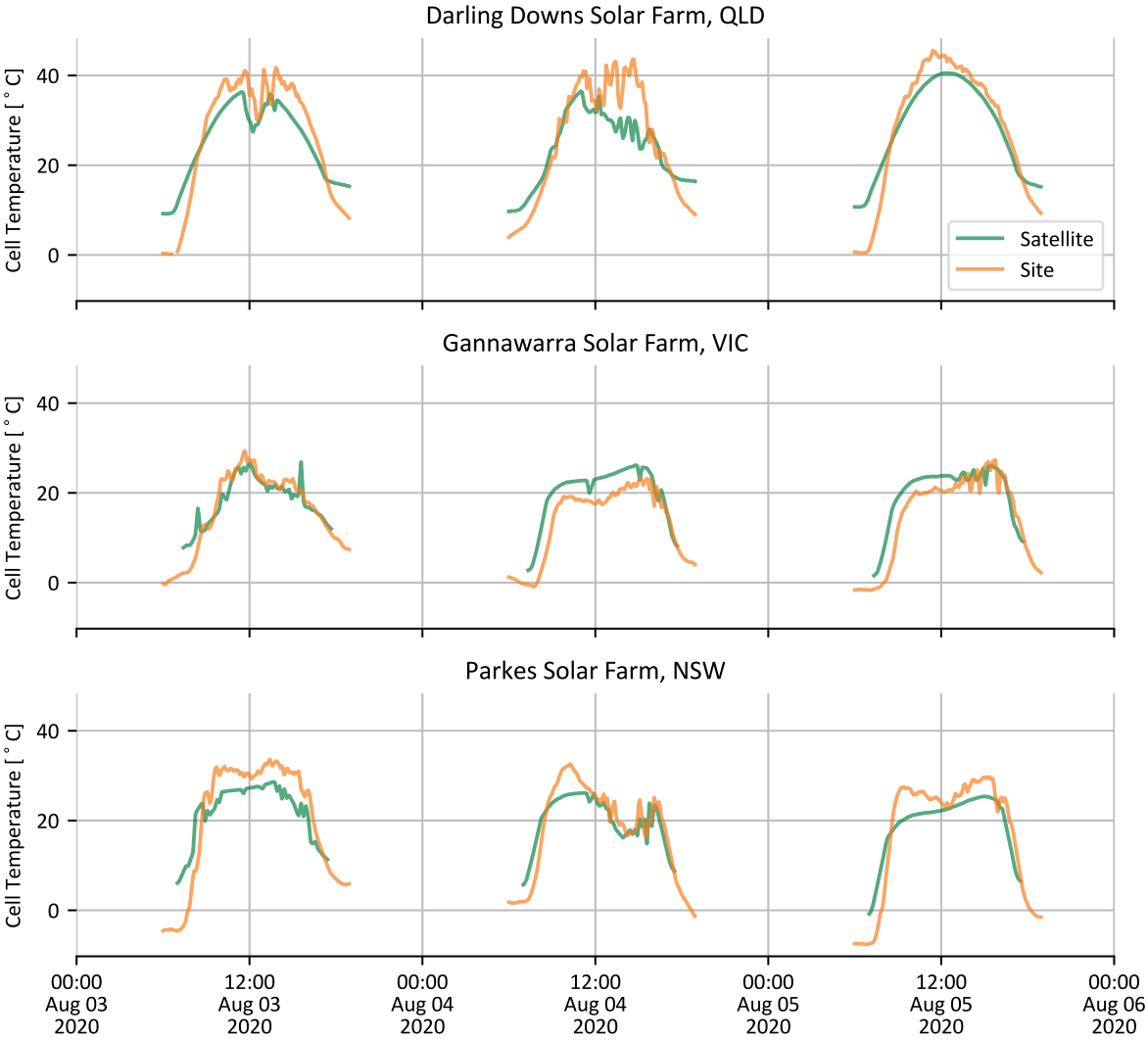


Figure 7. Comparing cell temperature from ground weather station and estimations derived from satellite data across three of ARENA's LSS projects over three winter days.

## QUANTIFYING UNCERTAINTY

Figure 8 shows the satellite bias with respect to site PV cell temperature measurements across 10 of ARENA's LSS projects over 5-minute and monthly timescales. Data in Table 5 provides MBE and RMSE values for satellite PV cell temperature measurements. The data shows that, on average, satellite cell temperature slightly underestimates site cell temperature measurements.

The relative MBE, standard deviation of MBE, and RMSE are all higher for cell temperature estimates than was observed for irradiance estimates. This likely reflects the fact that a single set of coefficients has been assumed to describe the cell temperature model at all sites. In reality, the way that weather conditions translate to module temperatures over the course of the day will vary depending on the particular module and installation conditions at each site. Future models could reduce the error associated with these estimates by conducting a correlation analysis similar to the 'Likely set-up' procedure described above for each site.

Similar to the observed effect for satellite measurements of irradiance, the spread and RMSE of the data reduces as the data is averaged at the monthly rather than 5-minute timescale.

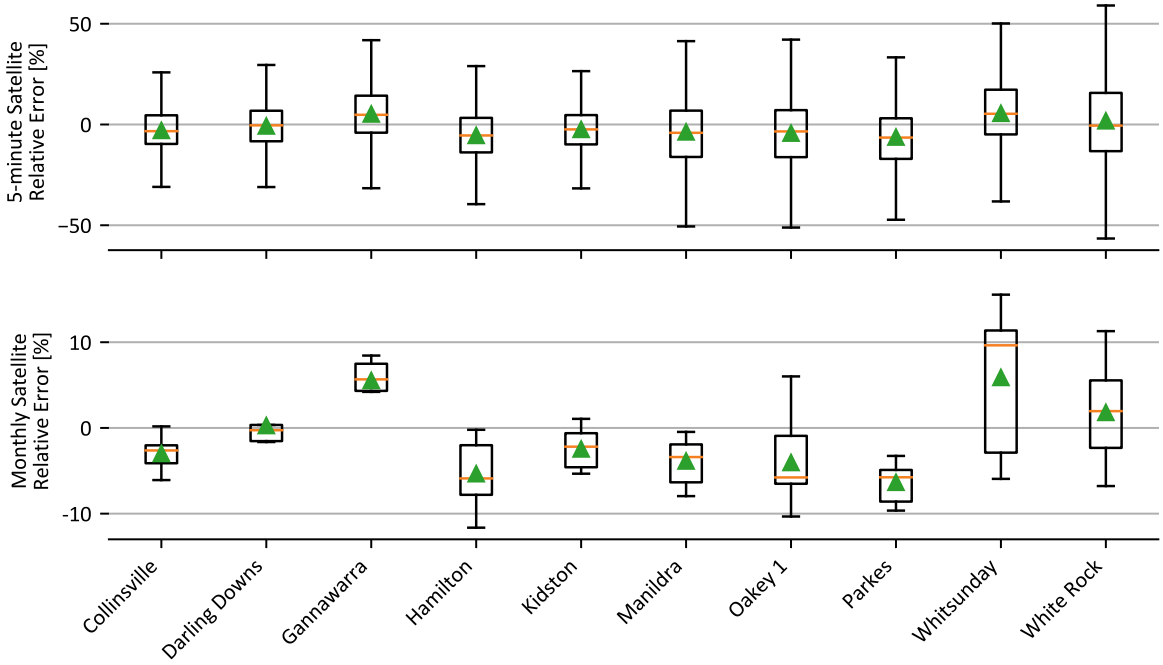


Figure 8. Describing the error between satellite and ground weather station PV cell temperature records at ten of ARENA's LSS projects (mean = green triangle, median = orange line). Monthly satellite bias error is the difference between satellite and ground weather monthly average temperatures. Error shown is as a percentage of the mean site value for the given time resolution.

TABLE 5. MBE AND RMSE FOR FIVE-MINUTE AND MONTHLY AVERAGE CELL TEMPERATURE ESTIMATIONS AT EACH PROJECT

| PROJECT | 5-MINUTE | | | | MONTHLY | | | |
|---|---|---|---|---|---|---|---|---|
| | MBE | | RMSE | | MBE | | RMSE | |
| | [°C] | [%] | [°C] | [%] | [°C] | [%] | [°C] | [%] |
| Collinsville | -1.10 | -2.99 | 4.23 | 11.55 | -1.09 | -2.99 | 1.28 | 3.51 |
| Darling Downs | -0.24 | -0.79 | 3.99 | 13.03 | 0.08 | 0.27 | 0.69 | 2.23 |
| Gannawarra | 1.48 | 5.17 | 4.54 | 15.82 | 1.53 | 5.53 | 1.67 | 6.04 |
| Hamilton | -2.22 | -5.47 | 6.26 | 15.47 | -2.15 | -5.35 | 2.52 | 6.27 |
| Kidston | -0.99 | -2.49 | 5.02 | 12.68 | -0.97 | -2.47 | 1.31 | 3.33 |
| Manildra | -1.13 | -3.73 | 6.31 | 20.83 | -1.14 | -3.87 | 1.34 | 4.56 |
| Oakey 1 | -1.50 | -4.34 | 6.78 | 19.58 | -1.38 | -4.05 | 2.06 | 6.05 |
| Parkes | -2.07 | -6.29 | 5.82 | 17.68 | -2.03 | -6.36 | 2.14 | 6.71 |
| Whitsunday | 1.99 | 5.50 | 5.88 | 16.20 | 2.11 | 5.88 | 3.33 | 9.27 |
| White Rock | 0.44 | 1.78 | 5.86 | 23.64 | 0.44 | 1.80 | 1.45 | 5.90 |
| **Average** | -0.53 | -1.37 | 5.47 | 16.65 | -0.46 | -1.16 | 1.78 | 5.39 |
| **Std. Dev.** | 1.36 | 3.99 | 0.91 | 3.65 | 1.38 | 4.13 | 0.72 | 1.93 |

The standard deviation of bias across all sites can be used to quantify the standard and expanded uncertainty for estimates for 5-minute PV cell temperature estimates. The temperature-corrected performance ratio calculation requires the PV cell temperature at each 5-minute interval to adjust the expected yield value to remove the effects of module temperature. A standard uncertainty of approx. ±4 per cent can be adopted. 5-minuite PV cell temperature estimates at other sites can be expected to have a mean bias error ±8 per cent of ground station measurements with 95 per cent confidence. Total uncertainty in the measurements can be estimated in the manner described for irradiance above. Module cell temperature sensors used for site measurements will typically have an expanded uncertainty of up to 2°C (or approx. 6 per cent based on the average cell temperature of 33.5°C across LSS sites) [19].[13]

13 In reality, the uncertainty of on-site measurements for module cell temperature will be lower as the average measurement from several reference cells are typically taken as the input for temperature-corrected performance ratio calculations.

The resulting total expanded uncertainty in the satellite-based 5-minute PV cell temperature estimates is ±3.4°C or 10 per cent at the 95 per cent confidence level. However, the level of uncertainty error introduced by the use of satellite data is of a similar magnitude to the uncertainty inherent in ground station measurements from typical sensors.

Furthermore, this uncertainty will have a limited impact on the uncertainty of the overall performance ratio calculation. Module temperature values are used to calculate an adjustment for cell temperature, equal to the difference in temperature from standard test conditions (25°C) multiplied by the module temperature coefficient of power. The module temperature coefficient of power is typically less than 0.5%/°C [20]. In this study, a value of 0.38%/°C has been assumed based on the average of the values from the 10 LSS projects used for validation. Using this value, the average module cell temperature across the LSS sites of 33.5°C would result in a temperature correction multiplier of 0.9677, i.e., the expected output at 33.5°C is 96.77 per cent of what would be expected at standard test conditions. An uncertainty of ±3.4°C corresponds to a temperature correction ranging from 0.9548 to 0.9806, i.e., ±1.3 per cent. This value can be used as the contribution of uncertainty in PV cell temperature measurements to the expanded uncertainty of the PR calculation [21]. The temperature correction is applied to irradiance at each 5-minute interval, which is then summed over the whole year. The relative error can be expected to remain approximately the same.

TABLE 6. UNCERTAINTY FOR 5-MINUTE SATELLITE-BASED 5-MINUTE CELL TEMPERATURE ESTIMATES

| | UNCERTAINTY RELATIVE TO GROUND STATION MEASUREMENTS | TOTAL UNCERTAINTY INCLUDING GROUND STATION INSTRUMENT UNCERTAINTY |
|---|---|---|
| **Standard Uncertainty** | ± 4.0% | ± 5.0% |
| **Expanded Uncertainty (95% confidence level)** | ± 8.0% | ± 10.0% |
| **Estimated Contribution to Expanded Uncertainty in PR (95% confidence level)** | ± 1.1% | ± 1.3% |

# ACCURACY OF SATELLITE PERFORMANCE RATIO ESTIMATES

The analysis of irradiance and cell temperature indicates that a reasonable degree of accuracy can be expected when estimating performance ratios using satellite-based data.

Three types of performance ratios are considered in this paper. These are explained in detail in the Definitions section, but are briefly described below.

- Raw Performance Ratio: This indicates overall performance of the solar farm, including the impacts of curtailment. No periods are excluded from the calculation except missing data. Based on generation data and annual irradiance only.
- Unconstrained Performance Ratio: This indicates the performance of farms excluding periods of curtailment and other constraints on farm output. Based on generation data and annual irradiance only. Performance ratio guarantees provided by EPC or O&M contractors typically use some kind of unconstrained performance ratio so that periods of underperformance due to network or other factors outside the control of the contractor are excluded from the calculation that determines whether liquidated damages are payable for underperformance.
- Temperature-Corrected Performance Ratio: This is the Unconstrained Performance Ratio corrected for losses due to module cell temperature, which enables comparison of farms in different regions and better estimates of relative performance at farms with several weeks of excluded data. Based on generation, irradiance and cell temperature data.

Generation data is metered and published by AEMO and provides the instantaneous output of the solar farm at each 5 minute period in the dataset. No uncertainty is associated with this measurement. Based on the formulas for Raw Performance Ratio and Unconstrained Performance Ratio, the uncertainty associated with estimating these from satellite data can be expected to be equal to the uncertainty associated with annual irradiance estimates. Note, the uncertainty associated with the detection and exclusion of constrained periods is not considered by this analysis.[14] Based on the formula for Temperature-Corrected Performance Ratio, the uncertainty associated with estimating this from satellite data can be expected to be equal to the combined uncertainty of annual irradiance and the contribution of cell temperature to performance ratio uncertainty. These uncertainties can be combined by summing in quadrature, assuming the sources of error are entirely independent. This can be considered an upper bound for the uncertainty, as there is a correlation between irradiance and cell temperature measurements.

The expected uncertainties associated with the different performance ratio estimates based on satellite data are shown in Table 7.

TABLE 7. EXPECTED UNCERTAINTY IN PR ESTIMATES RELATIVE TO VALUE OF ESTIMATE

| | RAW AND UNCONSTRAINED PR | | TEMPERATURE-CORRECTED PR | |
|---|---|---|---|---|
| | RELATIVE TO GROUND STATION MEASUREMENTS | TOTAL INCLUDING GROUND STATION INSTRUMENT UNCERTAINTY | RELATIVE TO GROUND STATION MEASUREMENTS | TOTAL INCLUDING GROUND STATION INSTRUMENT UNCERTAINTY |
| **Standard Uncertainty** | ± 2.3% | ± 2.5% | ± 2.5% | ± 2.8% |
| **Expanded Uncertainty (95% confidence level)** | ± 4.6% | ± 5.0% | ± 5.0% | ± 5.6% |

Note, in practice, errors associated with ground station instrument uncertainty are rarely considered when performance ratios are calculated for the purposes of monitoring solar farm performance and the associated contractual obligations. Provided that the instruments have been properly and recently calibrated and no errors are apparent on the face of the data, the instrument measurements are accepted as ground truth for subsequent performance ratio calculations. As such, for the remainder of this study, the reported uncertainties associated with calculated performance ratios only consider the uncertainty relative to equivalent measurements using ground station data.

14 As local limit data is available from AEMO from November 2021, any uncertainty associated with this calculation for future time periods will be greatly reduced, if not eliminated.

Figure 9 ranks the 10 LSS projects by the site temperature-corrected PR. The satellite estimates of raw, unconstrainted and temperature-corrected PR are shown alongside the corresponding values calculated using site data.

The lightest bars indicate that farms in Queensland (i.e., hotter temperatures on average) experience a more significant increase in performance ratio due to the temperature correction. However, the resulting temperature-corrected performance ratios present no strong pattern based on location, indicating the effectiveness of the correction in reducing differences in performance is solely attributable to environment.

Overall, there is a strong alignment between satellite and site values calculated at each of the sites. The data in Table 8 provides error values for satellite raw, unconstrained and temperature corrected PR estimates. These data show that, on average, satellite PR estimates slightly underestimate the raw PR, and slightly overestimate the unconstrained and temperature-corrected PRs relative to site PR values.

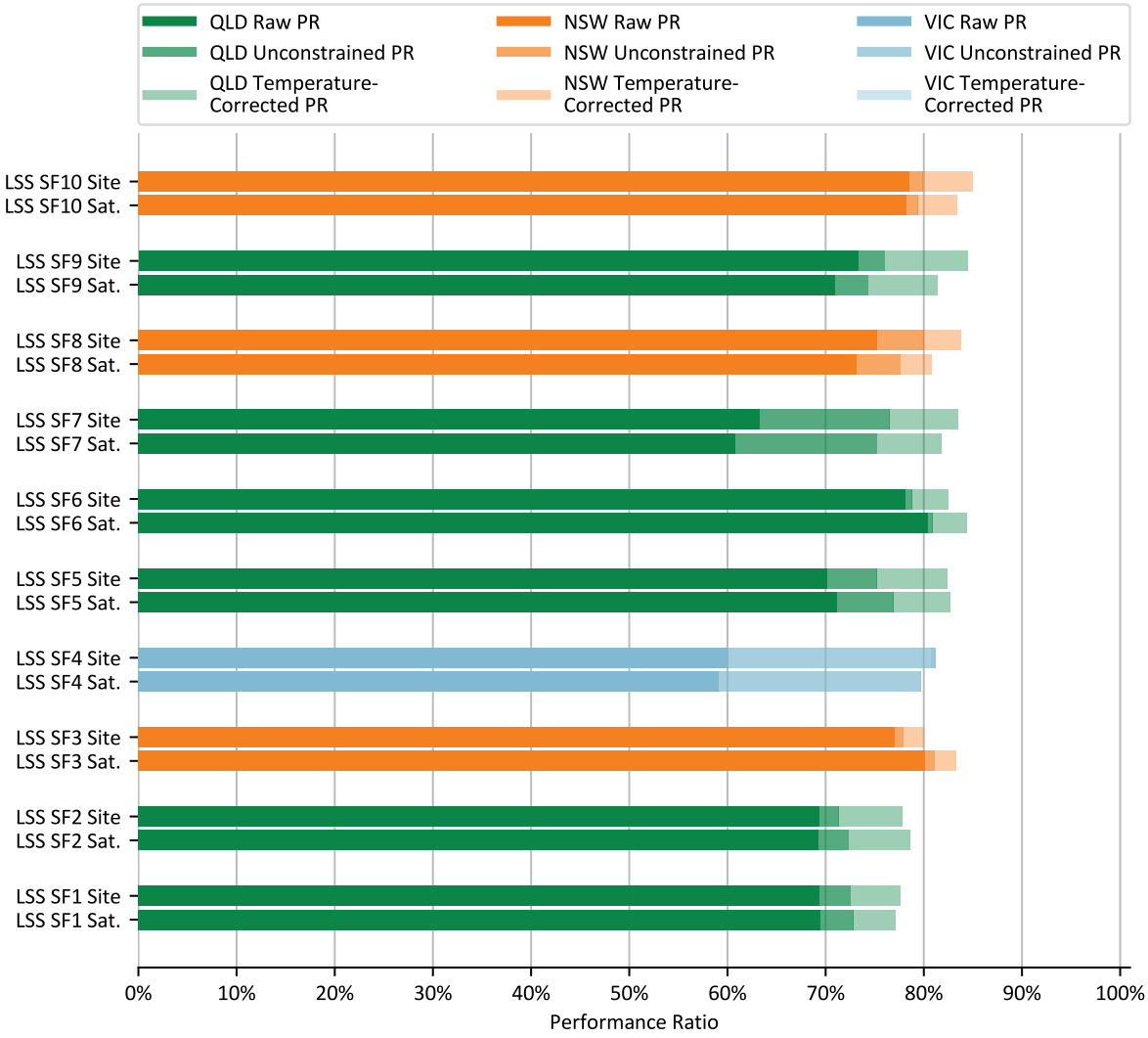


Figure 9. Comparing performance ratio estimates across ten of ARENA's LSS projects when using ground weather station and satellite irradiance data.

TABLE 8. MEAN BIAS ERROR FOR RAW AND WEATHER-CORRECTED ANNUAL PR ESTIMATES USING SATELLITE WEATHER DATA.

| PROJECT | ERROR [%] RAW PR | UNCONSTRAINED PR | WEATHER-CORRECTED UNCONSTRAINED PR |
|---|---|---|---|
| Darling Downs | 2.29 | 2.07 | 1.86 |
| Parkes | -0.24 | -0.49 | -1.62 |
| White Rock | 3.10 | 3.20 | 3.24 |
| Oakey 1 | 0.97 | 1.70 | 0.28 |
| Whitsunday | -2.46 | -1.23 | -1.74 |
| Hamilton | -2.30 | -1.73 | -3.14 |
| Manildra | -2.13 | -2.35 | -2.98 |
| Gannawarra | -0.97 | -1.47 | -1.09 |
| Kidston | -0.11 | 1.07 | 0.78 |
| Collinsville | 0.13 | 0.27 | -0.45 |
| **Average** | -0.17 (relative = -0.24%) | 0.10 (relative = 0.14%) | -0.49 (relative = -0.59%) |
| **Std. Dev.** | 1.80 (relative = 2.52%) | 1.76 (relative = 2.28%) | 1.95 (relative = 2.38%) |

As for irradiance and cell temperature, the standard deviation of the bias across all sites can be used to quantify the uncertainty associated with these estimates. The results indicate a very close alignment with the expected uncertainties from Table 7. From Table 8, a standard uncertainty (relative to the value measured) of ±2.5 per cent could be conservatively adopted for all three PR estimates. This is equal to the upper bound calculated in Table 7 for temperature-corrected PR based on the uncertainties in the input values.

Based on this standard uncertainty, PR estimates at other sites based on satellite data can be expected to fall within ±5.0 per cent of the estimated values with 95 per cent confidence. This means, for example, that if satellite-based estimates of temperature-corrected PR for a site is 81.8 per cent,[15] then there is a 95 per cent probability the actual performance as measured onsite falls between 77.7 per cent and 85.9 per cent, and it is more likely than not[16] that the site-measured PR falls between 79.8 per cent and 83.8 per cent [12]. While this degree of accuracy is not sufficient for the purposes of determining the liability of contractors under performance ratio guarantees, it is suitable for comparing the relative performance of different sites with a degree of confidence, for comparing the aggregate performance of multiple farms, and for setting a benchmark of observed performance across the industry as at 2020.

15 This is the average temperature-corrected PR at LSS sites in 2020.

16 To be precise, there is a 68% probability.

# PERFORMANCE AT SOLAR FARMS ON THE NEM

The comparison of ground station and satellite data at ARENA's LSS projects demonstrate a sufficient level of accuracy in satellite-based PR estimates for the purposes of comparing and benchmarking performance at large-scale solar farms across the NEM.

After filtering both satellite irradiance data and AEMO's public NEMWEB dispatch data in 2020 (calendar year), there are 26 large-scale solar farms with sufficient data to estimate all three performance ratios defined above. Figure 10 ranks the large-scale solar farms by temperature-corrected performance ratio. As expected, projects in Queensland experienced a larger temperature correction than any other state, as indicated by the lighter colour bars.

The data from Figure 10 is summarised by state in Table 9. The mean temperature-corrected performance ratio at farms in Queensland, NSW and Victoria were similar (78.3 per cent to 79.8 per cent). SA farms exhibited a substantially higher mean performance (83.0 per cent), though this was based on a sample size of 2. The mean performance ratio improved by over 10 per cent for farms in Victoria and SA, compared to less than 5 per cent for the average in Queensland and NSW. This likely reflects a higher level of curtailment occurring in the former states, as well as, particularly in the case of SA, smaller sample sizes.

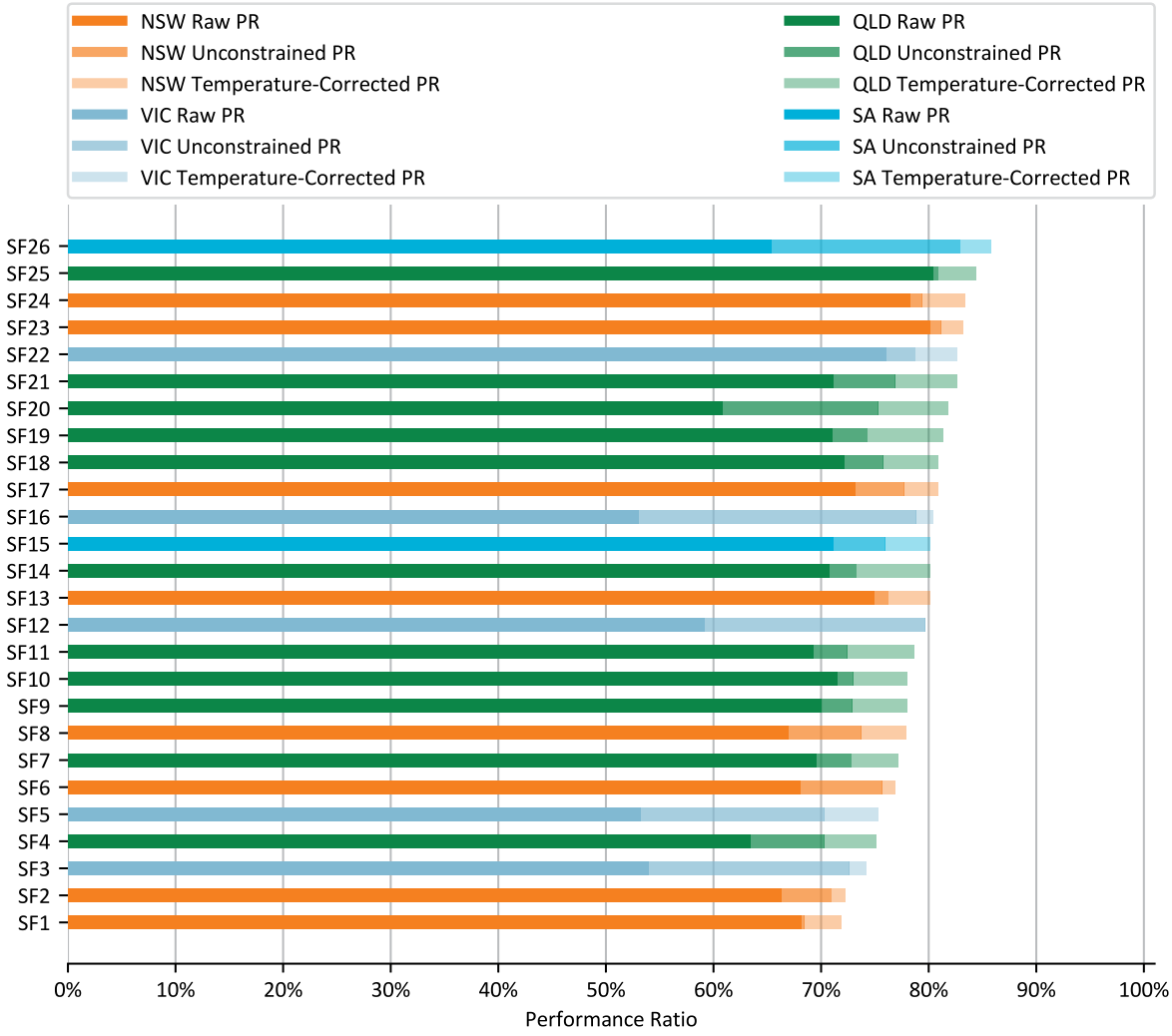

Figure 10. Performance ratio estimates for large-scale solar farms on the NEM using satellite weather data.

TABLE 9. SUMMARY OF KEY STATISTICS DESCRIBING PERFORMANCE RATIO ESTIMATIONS FOR LARGE-SCALE SOLAR FARMS ON THE NEM WHEN USING SATELLITE WEATHER DATA.

| | RAW [%] | UNCONSTRAINED [%] | TEMPERATURE CORRECTED [%] |
|---|---|---|---|
| QLD | 70.0 | 74.3 | 79.8 |
| NSW | 72.0 | 75.4 | 78.3 |
| VIC | 59.1 | 76.0 | 78.5 |
| SA | 68.3 | 79.5 | 83.0 |
| All | 68.4 (std = 7.53) | 75.4 (std = 3.75) | 79.3 (std = 3.58) |

Figure 11 presents boxplots of the 2020 raw, unconstrained and temperature-corrected PR values at the 26 solar farms in the sample. The raw PR values cover a much greater range than the unconstrained PR values. This is consistent with the large volume of curtailment observed at particular sites, especially in Victoria. The spread of the data reduced further after the temperature correction is applied, indicated both graphically in Figure 11 and by the reduction in the standard deviation in Table 9. This is consistent with the fact that the temperature correction has an unequal effect on sites depending on average local temperatures as noted above, hence the reduction in the difference between farms once this disparity is accounted for.

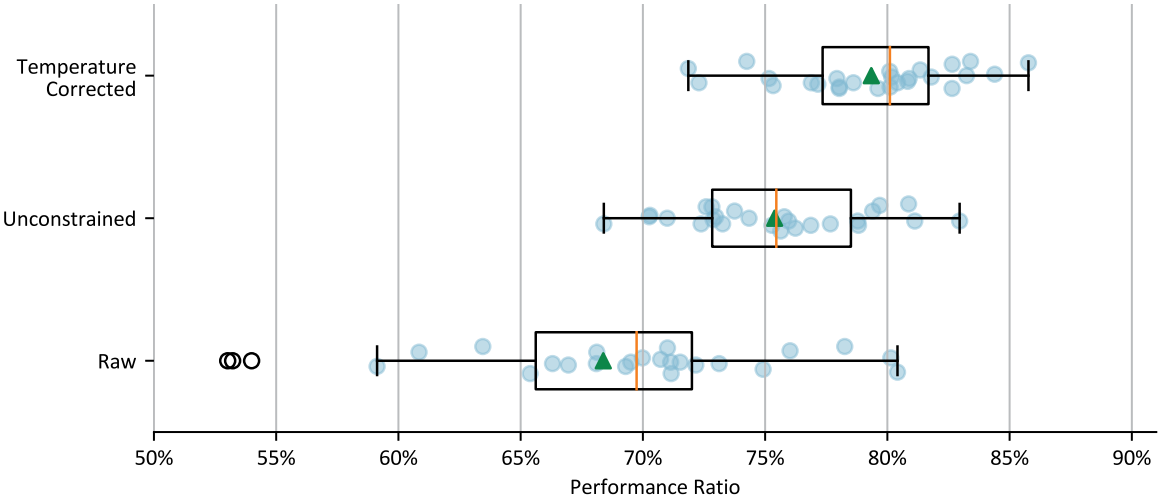


Figure 11. Boxplots comparing raw and temperature corrected PR for large-scale solar farms on the NEM (median = orange line, mean = green triangle).

The standard uncertainty of ± 2.5 per cent calculated in the previous section can now be applied to the complete dataset of 26 PR values to obtain confidence intervals on key metrics for the aggregate dataset. 95 per cent confidence intervals were generated by performing 10,000 Monte Carlo simulations for each of the three PR types. In each simulation, the PR value for each site was randomly drawn from a normal distribution with mean equal to the satellite PR estimate for the site and standard deviation equal to 2.5 per cent of the PR estimate for the site. The first quartile, mean and third quartile were calculated for each simulation. The confidence interval on the mean is based on the 0.025 and 0.975 quantiles for the means of the 10,000 simulations. The middle 50 per cent of solar farms are based on the 0.025 quantile of the first quartiles and the 0.975 quantile for the third quartiles of the 10,000 simulations.

TABLE 10. AVERAGE PERFORMANCE INCLUDING UNCERTAINTY FROM SATELLITE MEASUREMENTS

| | RAW PR [%] | UNCONSTRAINED PR [%] | TEMPERATURE CORRECTED PR [%] |
|---|---|---|---|
| **Mean** | 68.4 | 75.4 | 79.3 |
| **PR Range for Mean (95% Confidence)** | 67.7 - 69.0 | 74.7 - 76.1 | 78.6 - 80.1 |
| **PR Range for Middle 50% of Solar Farms (95% Confidence)** | 63.3 - 74.3 | 71.2 - 79.7 | 75.4 - 83.5 |

This data provides a useful benchmark for any industry stakeholders looking to understand where their projects sit with respect to performance of other large-scale solar farms. Owners or investors can determine whether the performance at a given solar farm is above or below average across the industry. Further, if a given solar farm's PR is below the lower limit of the middle 50 per cent range, it is likely to be in the bottom 25 per cent of farms ranked by performance.

The utility of this data as an industry benchmark is subject to several limitations, including:

- The sample size of just 26 farms over a single year in 2020 is relatively small.
- The methodology for identifying curtailment relies on visual data checks and a conservative approach that likely resulted in many periods of unconstrained operation being unnecessarily excluded from the calculation of unconstrained and temperature-corrected PRs.
- The many assumptions made about site setup noted throughout the paper, including azimuth, the initial fixed vs tracking approximation, the module temperature model, and module coefficient of power.
- Additional corrections sometimes made to performance ratio calculations, such as corrections for degradation or power factor, are not considered.

Nevertheless, the results remain valuable because:

- They represent the first NEM-wide assessment of performance at large-scale solar farms, building on the limited site-based PR results from LSS sites published in Report One from ARENA's Generator Operations Series: Large-scale Solar Operations (arena.gov.au/knowledge-bank/report-one-large-scale-solar-operations/).
- They demonstrate a robust methodology for satellite-based PR calculations that can be extended to additional farms and future years as more data becomes available.
- Although the results are from the single year 2020, the temperature-corrected performance ratio calculation results in a value that captures the underlying performance of the solar farms in question independent of the weather conditions in that year, and can thus be compared to performance ratios in future years.

# SUMMARY

Typically, high accuracy ground station irradiance and temperature measurements are used to measure the performance of a solar farm by calculating the performance ratio. The reliance on this data means that performance data for solar farms is temporally and spatially scarce, is expensive to maintain, and remains unavailable to the public. This paper explores the practicability of using satellite weather data and public electricity generation data to calculate PRs for large scale solar farms across the NEM.

Two variables required to calculate performance ratio in this study are effective irradiance and PV cell temperature. This study compares estimates of these variables derived from satellite data (satellite estimates) with actual measurements from ground-based weather stations (site data). This study demonstrates that while it is preferable to use ground-based weather data to perform PR calculations on large-scale solar farms, it is, within reason and depending on the application (e.g. broad understanding of performance as opposed to performance guarantee checks), possible to rely on satellite data, albeit introducing some additional uncertainty. Some applications, such as performance guarantees, require lower levels of uncertainty and therefore may require site data to be used.

Using satellite data, this paper introduces a solar PV performance benchmark across 26 Australian large-scale solar farms. Solar PV developers and investors can use this benchmark to broadly determine how the performance of their assets compare to others in Australia. This benchmark indicates the average PR for 26 Australian solar farms, after correcting for constrained periods and temperature, to be 79.3 per cent. The 95 per cent confidence interval on this value, accounting for the uncertainty introduced by using satellite data, is 78.6 per cent to 80.1 per cent. In addition, the performance ratio of the middle 50 per cent of solar farms in the sample fell between 75.4 per cent and 83.5 per cent, again with a 95 per cent confidence level. This data can be used to determine whether a given solar farm's performance is in the top or bottom 25 per cent of performance, relative to those sampled. This benchmark has been determined using satellite weather data purchased from Solcast and publicly available generation data from AEMO. Ground-based weather data and generation data was not used in creating this benchmark.

Improvement to the analysis is possible by reducing the omission of satisfactory periods of data. The present analysis method excluded a day of data if a local limit was visually determined to apply for any length of time on that day. AEMO have begun publicly recording local limit data tags on all generator assets in the NEM since November 2021. If this analysis were to be repeated on operational data post November 2021, then the unconstrained datasets will be more comprehensive with fewer unnecessary exclusions, improving the reliability of results.

Future study in this area should track how solar PV performance changes over time. The analysis in this paper is limited to the 2020 calendar year. The method used in this study to analyse solar PV performance could be extended across multiple years of data to track how changes in technology efficiencies and weather impact solar PV performance over time.

# REFERENCES


[1] NREL, "Weather-Corrected Performance Ratio," NREL, Albuquerque, 2013.

[2] NREL, "Solar Resource Glossary," NREL, [Online]. Available: https://www.nrel.gov/grid/solar-resource/solar-glossary.html#solarirradiance. [Accessed 21 April 2022].

[3] Sandia Labs, "PVPMC Home," Sandia Labs, [Online]. Available: https://pvpmc.sandia.gov/. [Accessed 5 May 2022].

[4] "Transpositional Model," PV Syst, [Online]. Available: https://www.pvsyst.com/help/models_meteo_transposition.htm. [Accessed 20 April 2022].

[5] Solcast, "Solcast," Solcast, [Online]. Available: https://solcast.com/. [Accessed 5 May 2022].

[6] Solcast, "Inputs and algorithms," Solcast, [Online]. Available: https://solcast.com/solar-radiation-data/inputs-and-algorithms/. [Accessed 5 May 2022].

[7] William F. Holmgren, Clifford W. Hansen, Mark A. Mikofski, "pvlib python: a python package for modeling solar energy systems," *Journal of Open Source Software,* vol. 3, no. 29, 2018.

[8] Sandia National Laboratories, "Photovoltaic Array Performance Model," Energy Systems Integration Group, 2022. [Online]. Available: https://www.esig.energy/wiki-main-page/photovoltaic-array-performance-model/. [Accessed 9 May 2022].

[9] GSES, "Effect of Orientation and Latitude on the Performance of Solar Tracking Arrays," [Online]. Available: https://www.gses.com.au/effect-of-orientation-and-latitude-on-the-performance-of-tracking-pv-arrays/. [Accessed 21 April 2022].

[10] Sandia Labs, "Sun Position," Sandia Labs, [Online]. Available: https://pvpmc.sandia.gov/modeling-steps/1-weather-design-inputs/sun-position/. [Accessed 5 May 2022].

[11] BrownDogGadgets, "Simple Dual Axis Solar Tracker," [Online]. Available: https://www.instructables.com/Simple-Dual-Axis-Solar-Tracker/. [Accessed 6 May 2022].

[12] Christian Reise, Björn Müller, David Moser, Giorgio Belluardo, Philip Ingenhoven, "Uncertainties in PV System Yield Predictions and Assessments," IEA-PVPS T13-12:2018, 2018.

[13] Manajit Sengupta, Aron Habte, Stefan Wilbert, Christian Gueymard, Jan Remund, "Best Practices Handbook for the Collection and Use of Solar Resource Data for Solar Energy Applications: Third Edition," NREL, Albuquerque, NM, 2021.

[14] T. C. Marcel Suri, "Satellite-based Solar Resource Data: Model Validation Statistics Versus User's Uncertainty," Presented at ASES SOLAR 2014 Conference, San Francisco, 7-9 July, 2014.

[15] Solargis, "Solar Resource Database: Validation of Solargis Solar Radiation Model," World Bank, Washington, DC, 2019.

[16] Aaron Habte, Manajit Sengupta, "Best Practices of Uncertainty Estimation for the National Solar Radiation Database (NSRDB 1998-2018)," NREL Conference Paper NREL/CP-5D00-70165, 2017.

[17] D. King, "Sandia Photovoltaic Array Performance Model," Sandia National Laboratories, Albuquerque, NM, 2004.

[18] pvlib, "Source code for pvlib.temperature," pvlib, [Online]. Available: https://pvlib-python.readthedocs.io/en/stable/_modules/pvlib/temperature.html. [Accessed 21 April 2022].

[19] "International Standard IEC 61724-1:2017," 2017.

[20] HOMER Energy, "HOMER Pro Version 3.7 User Manual," HOMER Energy, Boulder, 2016.

[21] K. Emery, "Uncertainty Analysis of Certified Photovoltaic Measurements at the National Renewable Energy Laboratory," NREL, 2009.

[22] NREL, "Weather-Corrected Performance Ratio," NREL, Albuquerque, 2013.

[23] NREL, "Analysis of Photovoltaic System Energy Performance Evaluation Method," NREL, Albuquerque, 2013.

[24] Sandia National Labs, "PVPerformance Modeling Collaborative," Sandia National Labs, [Online]. Available: https://pvpmc.sandia.gov/. [Accessed 5 March 2022].

[25] ARENA, "Large-scale solar operations," ARENA, Canberra, 2021.

[26] ARENA, "Negative pricing and bidding behaviour in the NEM," ARENA, Canberra, 2021.

[27] "Solar Irradiance," NASA , 01 January 2008. [Online]. Available: https://www.nasa.gov/mission_pages/sdo/science/Solar%20Irradiance.html. [Accessed 20 April 2022].

[28] NASA, "Brian Dunbar," NASA, 01 January 2008. [Online]. Available: https://www.nasa.gov/mission_pages/sdo/science/Solar%20Irradiance.html. [Accessed 20 April 2022].

[29] Kevin Anderson, Mark Mikofski, "Slope-Aware Backtracking for Single-Axis," NREL, Albuquerque, NM, 2020.

# DEFINITIONS

TABLE 11. DEFINED TERMS

| TERM | DESCRIPTION |
|---|---|
| Performance Ratio | **Performance Ratio (PR)** is a measure of overall plant efficiency. It is calculated as the ratio of the energy generated with respect to the energy that would have been generated if the system operated continuously at the rated efficiency of the modules at nominal standard test conditions (STC). The PR measures the impact of all losses experienced by the system from the sunlight reaching the modules to energy export to the grid, including module degradation, reflectance, temperature losses, inverter efficiency, and all other balance of system component inefficiencies [22]. Engineering, procurement, and construction (EPC) contracts for PV projects typically stipulate liquidated damages based on a PR guarantee value, with the plant being expected to perform at or above this PR value over a pre-agreed period of operation following practical completion. The PR guarantee is then compared to actual plant PR to determine any recoverable damages. As there are many external and uncontrollable factors that impact PR (including, for example, where a plant is required to reduce its generation for grid stability reasons), the operational data that is included and/or excluded from the PR assessment is often determined on a bespoke basis [23].<br>Under the terms of the contract, the EPC contractor may be required to pay performance liquidated damages (PLDs) equivalent to the value of the energy forgone by the owner due to the failure to achieve the agreed plant performance. If a performance shortfall is unable to be rectified, the EPC contractor may be required to pay lifetime PLDs equivalent to the total value of all future foregone revenue. This study does not break down the PR at this level of detail, nor determine whether poor performance is attributable to the EPC contractors.<br>The performance ratio can be calculated in different ways depending on the purpose. The formulas for the Raw Performance Ratio, Unconstrained Performance Ratio and the Temperature-Corrected Performance Ratio are shown in the rows below. |
| Raw Performance Ratio | $$PR = \frac{Y_f}{Y_{ref}} = 100 \frac{\sum_k P_{out,k} \times \tau_k}{\sum_k P_{STC} \left( \frac{G_{POA,k}}{G_{STC}} \right) \times \tau_k}$$<br>Where:<br>- $Y_f$: Actual Energy Yield over the measurement period [MWh]<br>- $Y_{ref}$: Reference Energy Yield over the measurement period [MWh]<br>- $PR$: Raw Performance Ratio [%]<br>- $P_{out,k}$: Measured AC power output of the solar farm at interval $k$ [MW]<br>- $P_{STC}$: Array DC power rating; the total DC power output of all installed PV modules at STC [MW]<br>- $G_{POA,k}$: Measured plane of array (POA) irradiance at interval $k$ [W/m$^2$]<br>- $G_{STC}$: Irradiance at standard test conditions [1,000 W/m$^2$]<br>- $k$: A given recording interval<br>- $\tau_k$: The length of the recording interval $k$; in this study 5-minute intervals have been used [h] |
| Unconstrained Performance Ratio | The unconstrained performance ratio is calculated in the same manner as the Raw Performance ratio, however, the following periods are excluded:<br>- Night-time periods: times when the clear-sky global horizontal irradiance was equal to zero.<br>- Negative pricing: where the 5-minute spot price or 30-minute settlement price were negative.<br>- Semi-dispatch cap: where the semi-dispatch cap was active requiring the solar farm to follow dispatch signals.<br>- Local limit active at solar farm (this was manually estimated by visualising data) |

| TERM | DESCRIPTION |
|---|---|
| Temperature-Corrected Performance Ratio | $$PR_{temp\_corr} = 100 \frac{\sum_k P_{out,k} \times \tau_k}{\sum_k P_{STC} \left(\frac{G_{POA,k}}{G_{STC}}\right) \times \left(1 + \frac{\delta}{100}(T_{mod,k} - T_{mod,ref})\right) \times \tau_k}$$<br>Where:<br>- $PR_{temp\_corr}$: Temperature-corrected performance ratio [%]<br>- $T_{mod,k}$: Average measured module temperature at interval $k$ [°C]<br>- $T_{mod,ref}$: Constant reference module temperature; in this study $T_{mod,ref} = T_{STC} = 25$ [°C]<br>- δ: Temperature coefficient of power corresponding to installed modules [%/°C, negative in sign]<br>- See Raw Performance Ratio for other parameter definitions.<br>This was calculated for the same periods as the Unconstrained Performance Ratio. This represents the performance of the plant if losses due to module temperatures above or below STC and any constraints are ignored. |
| pvlib | pvlib python is a community supported tool that provides a set of functions and classes for simulating the performance of photovoltaic energy systems. pvlib python was originally ported from the pvlib MATLAB toolbox developed at Sandia National Laboratories and it implements many of the models and methods developed at the Labs. More information on Sandia Labs PV performance modeling programs can be found at the PV Performance Modeling Collaborative [7, 24]. |
| The Perez Model | This study uses the Perez effective irradiance model, which estimates the total sky diffuse irradiance received by a module tilted from the horizontal. Note that the albedo is not considered but maybe added to obtain total hemispheric diffuse radiation on a tilted module. The output has been shown to be suitable for most solar gain calculations. The model represents the sky dome as an isotropic background where circumsolar and horizon effects are superimposed. The circumsolar effect is simulated by a point source at the sun's position, and the horizon effects by a linear source at the horizon, which can either be a positive or negative source signifying horizon and zenith brightening, respectively.<br>pvlib takes extra-terrestrial normal irradiance, direct horizontal irradiance (DHI), direct normal irradiance (DNI), solar zenith angle, solar azimuth angle and the relative airmass value to calculate the solar zenith angle, sky's clearness, sky's brightness and atmospheric precipitable water content. These results are then applied to the diffuse irradiance on a tilted module which calculates the variable as a function of DHI, module's slope, and two coefficients F1 and F2 expressing the degree of circumsolar and horizon anisotropy, respectively. The coefficients F1 and F2 vary according to the sky's clearness and are split into eight separate bins ranging from overcast to clear sky conditions. Additionally, there are two terms which are a function of the incidence angle of the sun on the module and the solar zenith angle, respectively. |
| Mean Bias Error (MBE) | $$MBE = \frac{\sum_{i=1}^{n} error_i}{n}$$<br>Where:<br>$$error_i = y_{model,i} - y_{actual,i}$$<br>Indicates whether a model consistently overestimates or underestimates the target value. |
| Standard Deviation of Bias (SD) | $$SD = \sqrt{\frac{\sum_{i=1}^{n}(error_i - \overline{error})^2}{n}}$$<br>Indicates the variation in the bias over the modelled values. |
| Root Mean Square Error (RMSE) | $$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(error_i)^2}{n}} = \sqrt{MBE^2 + SD^2}$$<br>Measures the average error of model without considering error direction and gives a relatively high weight to large errors. Reflects both the bias and variation in the error of the data. |
| Standard Uncertainty (u) | The standard uncertainty associated with an estimate of a quantity, equal to one standard deviation. This can be expressed as an absolute value or as a percentage of the mean value of the quantity being estimated. |
| Expanded Uncertainty (U) | $$U=ku$$<br>The uncertainty associated with an estimate at a given confidence level. The standard uncertainty can be converted to an expanded uncertainty using a coverage factor, k. Where the uncertainty follows a normal distribution, the coverage factor is approximately k = 2 for a 95% confidence level. |

Further information is available at
arena.gov.au

**Australian Renewable Energy Agency**

**Knowledge Sharing Team**
knowledge@arena.gov.au

*Postal Address*
GPO Box 643
Canberra ACT 2601

*Location*
2 Phillip Law Street
New Acton ACT 2601

*Engage with us*

ARENAWIRE



Australian Government
Australian Renewable Energy Agency

ARENA